

EDUCATION POLICY ANALYSIS ARCHIVES

A peer-reviewed scholarly journal

Editor: Sherman Dorn

College of Education

University of South Florida

Copyright is retained by the first or sole author, who grants right of first publication to the **Education Policy Analysis Archives**. EPAA is published jointly by the Colleges of Education at Arizona State University and the University of South Florida. Articles are indexed in the Directory of Open Access Journals (www.doaj.org).

Volume 13 Number 32

June 15, 2005

ISSN 1068-2341

Academic Freedom, Tenure, and Student Evaluations of Faculty: A Response to Haskell and His Critics

Edith J. Cisneros-Cohernour
Universidad Autónoma de Yucatán
Mérida, Yucatán, México

Citation: Cisneros-Cohernour, E. J. (2005, June 15). Academic Freedom, Tenure, and Student Evaluations of Faculty: A Response to Haskell and His Critics. *Education Policy Analysis Archives*, 13(32). Retrieved [date] from <http://epaa.asu.edu/epaa/v13n32>.

Abstract

I comment on the strengths and limitations of Haskell's article and provide a critical review of his arguments about the negative impact of SEF on tenure and other administrative decisions. I object to the limited evidence supporting the claim that the use of student evaluations per se challenges academic freedom.

This article is in response to Haskell's "Academic Freedom, Tenure & Student Evaluations of Faculty: Galloping Polls In the 21st Century" (1997) published in this journal as issue Number 6 of Volume 5, in which it is stated that student evaluations of faculty (SEF) "impinge on academic freedom" and tenure. Haskell combines arguments and data, and explores implications of SEF covered throughout the literature to build his case against the use of student evaluation results. Haskell's subsequent articles have focused on views of the court concerning the use of student ratings for promotion, reappointment, and tenure, and the analysis and implications of court views of SEF instruments' accuracy and psychometric validity, as well as teaching quality (Haskell, 1997a, 1997b, 1997c, 1997d).

I agree with Haskell on some issues. I agree that Haskell's claim that SEF infringe on faculty's academic freedom is a "novel" and thought provoking view of the possible consequences of SEF not thoroughly discussed in the literature. I also share Haskell's

concerns about possible misuse of SEF for tenure and other administrative decisions, and believe that more attention needs to be given to validity issues. However, more empirical evidence is needed in order to support Haskell's view that SEF challenges academic freedom.

Strengths

Haskell discusses various important issues that deserve special recognition. First, he focuses on the possible impact of SEF on professors' academic freedom, which is a novel idea not fully studied in the literature on faculty evaluation. The paper reminds the reader of the complexities of evaluating teaching by including the views of current research on SEF and other aspects rarely addressed in the literature, as well as information on administrative pressures and the legal and political aspects of the evaluation.

Second, Haskell's paper raises important concerns regarding the consequences of evaluation, especially the misuse of evaluation results and the legal consequences that could result. Legal considerations are one of Haskell's strongest points. Misusing the ratings and releasing them to the public could lead to serious legal consequences affecting a professor's reputation and professional life. The paper also emphasizes the need for readers to reflect on how the evaluation system may impact classroom practice for faculty, students and the institution as a whole.

Haskell's argument about the validity of the ratings is another strong point of the article. Validity is an issue in the evaluation of teaching that has already been widely discussed in the literature. Most of the researchers on college teaching believe the ratings are valid and reliable because various studies have found little variance across responses (Marsh, 1987) and because when SEF has been followed over time, student responses have not significantly changed (Frey, 1976). The validity of the ratings has also been examined by Marsh and Bailey (1993), who believe that SEF are relatively valid against a variety of indicators of effective teaching. In a review of the research on SEF, El-Hassan (1995)¹ found that in general the literature supports a modest degree of validity of the ratings. Although such arguments support the statistical validity of student ratings, they say little about their consequential validity. The problems of construct validity and the issues of possible untrained raters and untrained decision-makers could lead to invalidity and misuse of rating results. Most of the research has focused on the possible factors that could affect the ratings, but if SEF results are misinterpreted and misused, then consequential validity becomes compromised.

Haskell's article illustrates how the recommendations on the use of the ratings made by the main researchers in the field are not taking place in practice. Various researchers in the evaluation of teaching, such as Braskamp and Ory (1994), Centra (1993), Marsh (1984, 1982, 1980) and Marsh and Dunkin (1992), Menges and Brinko (1986) and others, have written recommendations and limitations when using and interpreting SEF results. Some of the recommendations refer to: (1) the use of the ratings for formative purposes; (2) the limitations of SEF and the suggestions for their interpretation; (3) the advised use of the ratings with other sources, such as peers, external observers, and self-evaluation. These

¹ El-Hassan (1995) provides an overview of the literature on student ratings of instruction.

researchers have also discussed the importance of considering the type of course when analyzing and interpreting SEF results. Unfortunately, as Haskell's cases illustrate, reality ignores these recommendations. The data from the evaluation are being used for making administrative decisions, SEF results are used as the only source for evaluating teaching, and although decision-makers receive information on the type of course, they compare professors' ratings without considering these and other important factors. The consequences of ignoring this reality are serious and more research is needed on the effects of the current star system. In addition, the possible impact of SEF on academic freedom is an important issue that deserves further consideration. This claim, however, needs to be proven.

Since universities are facing public pressures for accountability, it is not strange to find that they are paying attention to SEF as an indicator of institutional effectiveness. Keeping students as a source for evaluating college teaching may respond not only to democratic principles but also to a political process. Student involvement in the evaluation is important but problematic when non-academic political constraints cause faculty members to change their focus from student learning to an ill-fated focus on earning high SEF ratings. Haskell believes that SEF is leading professors to improve their ratings even if this means to inflate the grades of their students. The literature on college teaching is divided on the topic of SEF impact on grade inflation, but recent articles have provided some support for Haskell's views.

Limitations

Haskell's paper presents two main weaknesses not fully addressed by his critics: lack of evidence to support his arguments that SEF challenges academic freedom, and problems with his interpretation of research findings. In addition, his decision to use an extended definition of academic freedom brings up new issues related to freedom of speech and its possible abuses.

Problems with evidence

The lack of sufficient evidence is clear when Haskell cites Dershowitz to support his point that SEF impinges on academic freedom. Professor Dershowitz is a tenured professor who was not affected by the remarks that he made in his classroom. Although some of his students gave him low ratings for his teaching, he continues teaching what he believes and did not suffer pressures from the administration to change his teaching or limit his course coverage. So, one can not argue that the academic freedom of the instructor was limited because of SEF. There is a possibility, as Dershowitz himself noted (1994), that an untenured professor will be under pressure to change his teaching in order to satisfy student requests and raise his SEF results. If Haskell had provided a case in which this actually happened, then he would have support for his argument, but he failed to provide such a case. The Dershowitz case, as I discuss later, is also problematic as a source of evidence because it raises issues about freedom of speech.

Lack of evidence is also present in the three cases that Haskell uses to illustrate how SEF infringes on classroom curriculum. The cases, although important, do not clearly show how SEF is the cause for curriculum changes. In case one, it is unclear if the instructor was dismissed because of student pressures or if there was a legitimate reason for dismissal based

on teaching incompetence. Since Haskell does not provide additional information about the case, his example is not useful to support his argument. In case two, he used an indirect source to support his views that SEF impinges on academic freedom. In this case, a single faculty member affirms knowledge of six other professors whose academic freedom has been affected. Since the professor did not suffer limits on his own academic freedom, and it is unclear if he witnessed how his colleagues' freedom was affected, this could constitute simple "hearsay" for legal purposes. Case three shows how administrators reassigned an instructor who obtained low rating results to other courses in which she could obtain higher results. The case is more useful to illustrate the importance of considering the type of course when interpreting SEF results (e.g., required versus elective courses), and how administrators use the system and help their faculty to improve their ratings. Haskell avoided criticism by indicating that he uses the cases as "illustrations of the role of SEF, making no attempt to assess validity or non validity of the examples". However, this does not excuse him from providing valid evidence to make his case against the use of SEF for administrative decisions. By providing such weak evidence he puts himself in a situation similar to that in the studies that he criticizes.

In another part of his article, Haskell states that when students lack the required knowledge, instructors have to lower academic standards and adapt the course to student needs. Although this is an important issue, Haskell needs to show how the lack of students' knowledge background could affect the evaluation and have consequences for professors' academic freedom. Adapting the course to student needs or covering the whole program to provide required knowledge for future courses without considering the students constitutes a dilemma for many instructors, but this does not necessarily show infringement on academic freedom. Haskell needs to provide at least one supporting case.

Problems with interpretation

There are also various problems with the interpretation of research findings in Haskell's article. First, Haskell states that if a professor alters his/her behavior and then improves his/her ratings, this is evidence of how SEF is leading professors to manipulate the ratings and limit their own academic freedom. However, the circumstance that a professor changes his/her behavior and then there is a change in his SEF results is not necessarily a causal relationship. These findings can also be interpreted in the opposite way: if the professor improves his/her teaching, the ratings may improve because he is becoming a better instructor. The relationship between ratings and faculty behavior can take place either way.

Second, Haskell cites one survey in which one-third of the faculty perceived SEF to be a challenge to their academic freedom. However, it is unclear what the perceptions are of the other two-thirds of the subjects who participated in that survey. In addition, another problem of interpretation occurs when Haskell states that faculty have been dismissed because of their low SEF results. Although this may be the apparent cause of dismissal, there is a possibility that the professors were dismissed because of other reasons. After all, professors are hired not only for working in the classroom but to conduct research and to serve on committees as participants in their organizations. A similar problem exists when Haskell states that SEF shape faculty behavior to avoid dismissal. This may be true in some cases but not necessarily in all. He also implies that students would only enjoy courses that

are graded "easy," but this excludes the possibility that students could certainly enroll in a course for other reasons.

Haskell believes that the public release of SEF ratings is inappropriate, but he neglects addressing two possibilities: first, the institution publishes SEF results for all faculty including those who obtained low ratings, and second, the institution publishes only SEF results for those who obtained higher ratings. Has the academic freedom of those who obtained higher SEF results been impinged on by publishing their names in a list of excellent instructors? When invalid data are available to the public about the professor's low SEF results, there could be a possibility of defamation, but when the institution publishes only information about those who obtain higher ratings, does it constitute an infringement of the professor's academic freedom?

Definition of Academic Freedom

Haskell's decision to support an extended definition of academic freedom is problematic because it could lead to the same negative consequences of limiting access to higher education by women and minorities. Although he refers to academic freedom as a right that allows professors to pursue a line of inquiry or teaching without being "penalized, censored or fired", he agrees with Dershowitz's statement that professors should be protected against pressures "to censure unpopular subjects within the popular belief systems of students." The problem with the last statement is that this broad interpretation of academic freedom could lead to negative consequences if the instructor invokes his right to academic freedom when creating a threatening or harassing environment for women and minorities. There is a thin line between freedom of speech and its abuse. More information is needed about this case in order to know if Dershowitz used his right to academic freedom simply to justify a limited coverage of rape or the creation of a hostile environment for women in the class.

In his *EPAA* commentary on Haskell's article, J. Stake (1997) asserted that academic freedom is limited, since professors are not "free from constraints" imposed by the academy or by university regulations of research. I share Stake's perception that academic freedom is not an unrestricted right of instructors, but I believe that academic freedom belongs to the individual faculty member, not the academy of professors or the administration. I also believe that academic freedom is a limited right because in a democratic society individual rights are constrained by the rights of others. I do agree with Stake that it would be ideal if the violations of academic freedom could be resolved within the academy rather than in the courts, but I would not like the university to legislate on individual rights. The protection of individual and society rights goes beyond the university arena.

Haskell's article has also received commentary from Theall (1997). Theall expressed his disapproval of Haskell's not providing constructive criticism, including a limited examination of the research on student ratings, and misinterpreting Theall's comments on the use of SEF results. Although Haskell recognizes that there are faculty who believe that SEF does not influence their behavior and therefore, do not see SEF as an "abridgment of academic freedom", I agree with Theall that Haskell has provided more evidence to support his views than to contradict them. It is evident that Haskell is writing about an issue that is important to him, not only as a researcher but also as an instructor. I agree that more constructive

criticism would enhance the argument, but I think that the bias present in Haskell's work is also present in other studies where the researchers say little or nothing of their dual roles as researchers and administrators. In addition, I agree with Theall's perception that Haskell does not provide an in-depth analysis of the wide body of research on student ratings. I, however, think that Haskell's omission is not a serious mistake since the purpose of the article is not to question the validity of the research on student ratings, but to criticize the use of the ratings for tenure and other administrative decisions². Even though Haskell has not altered Theall's quotation on the limitations of the ratings, I agree with Theall that Theall's quotation is used more as an argument against the use of SEF than as a recommendation to improve the practice of using the ratings.

What I do not share with Theall is the opinion that the literature on student ratings should not be recommended to those who are not "actively involved in ratings research or practice because such writings can mislead readers who aren't really familiar with the cited ratings literature." I do not agree with this because I believe that research findings can not be exclusive only for an elite group of researchers. Such a limited audience will deprive scholarship of valuable multiple perspectives and is contrary to the nature of responsible scientific inquiry. Haskell's work has weaknesses as well as strengths. Even though he omits important studies in his discussion of the use of student ratings, lacks enough evidence to support some of his arguments, and has problems with interpretation, the article provides an opportunity to reflect on the use of evaluation results and their possible consequences. Evaluators should clearly benefit from keeping in mind the importance of the evaluation context, and the legal and ethical consequences of their work.

Overall, Haskell has made a strong argument against the use and misuse of student ratings in tenure and other administrative decisions. His concerns on this point are also shared by other researchers in the field, including those who support the validity of the ratings. The weakness in his argument lies in lack of evidence to prove that SEF violates the academic freedom of the professors.

Conclusion

The evaluation of teaching is complex. Haskell's work is a reminder of this complexity because it brings up important points that deserve attention for those interested in the evaluation and improvement of college teaching. Some of Haskell's ideas are not new. Several of his arguments about the limitations and problems of SEF have been stressed by various researchers in the field (e.g., Braskamp & Ory, 1994; Centra, 1993; Theall & Franklin, 1991, and others). His argument also presents weaknesses that deserve attention and require more evidence to prove that SEF violates professorial academic freedom. What is new and deserves praise is that he reminds the reader that the evaluation of college teaching is not only the domain of researchers and scientific inquiry, but also an important area for teachers' unions, administrators, lawyers, politicians, and the public outside academia.

² "It is important to note at the outset, that it is not SEF per se that is the issue, but the impact of its use on salary, promotion, tenure decisions, and equally important, its impact on the delivery of quality education." (Haskell, 1997)

Evaluators and researchers can benefit from the reminder that teaching assessment takes place in a context in which multiple factors operate, and in which all those involved, centrally and peripherally, have a stake.

References

- Benson, D. E. Lewis, J. M. (1994). Students' evaluation of teaching and accountability: implications from the Boyer and the ASA reports. *Teaching Sociology*, 22, 195-99.
- Braskamp, L. A., & Ory, J. C. (1994). *Assessing faculty work: Enhancing individual and institutional performance*. San Francisco, CA: Jossey Bass.
- Centra, J. A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco, CA: Jossey-Bass Inc.
- Dershowitz, A. (1994). *Contrary to popular opinion*. New York: Berkeley Books.
- Dilts, D. A., Samavati, H., Moghadam, M.R., and Haber, L.J. (1994). Student evaluation of instruction: Objective evidence and decision making. *Journal of individual employment rights*, 2, 73- 86.
- El-Hassan, K (1995). Students' Ratings of Instruction: Generalizability of Findings. *Studies in Educational Evaluation*, 21, (4), 411-29.
- Frey, P. (1976) Validity of student instructional ratings. Does timing matter? *Journal of Higher Education*, 47 (3), 327-336.
- Haskell, (1997a). Academic freedom, tenure, and students evaluation of faculty: Galloping polls in the 21st Century. *Education Policy Analysis Archives*, 5(6). Retrieved June 15, 2005 from <http://epaa.asu.edu/epaa/v5n6.html>
- Haskell, (1997b). Academic freedom, promotion, reappointment, tenure, and the administrative use of student evaluation of faculty: (Part II) views from the court. *Education Policy Analysis Archives*, 5(17). Retrieved June 15, 2005 from <http://epaa.asu.edu/epaa/v5n17.html>
- Haskell, (1997c). Academic freedom, promotion, reappointment, tenure, and the administrative use of student evaluation of faculty (SEF): (Part III) Analysis and implications of views from the court in relation to accuracy and psychometric validity. *Education Policy Analysis Archives*, 5(18). Retrieved June 15, 2005 from <http://epaa.asu.edu/epaa/v5n18.html>
- Haskell, (1997d). Academic freedom, promotion, reappointment, tenure, and the administrative use of student evaluation of faculty (SEF): (Part IV) analysis and implications of the court in relation to academic freedom, standards and quality of instruction. *Education Policy Analysis Archives*, 5(21). Retrieved June 15, 2005 from <http://epaa.asu.edu/epaa/v5n21.html>
- Marsh, H. A. (1987) Students' evaluations of university teaching: research findings,

methodological issues, and directions for future research. *International Journal of Educational Research*, 11, 253-388.

Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76, 707-754.

Marsh, H. W. (1982). Validity of students' evaluations of college teaching: A multitrait-multimethod analysis. *Journal of Educational Psychology*, 74, 264-279.

Marsh, H. W. (1980). The influence of student, course, and instructor characteristics on evaluations of university teaching. *American Educational Research Journal*, 17, 219-237.

Marsh, H. W. and Bailey, M. (1993). Multidimensional students' evaluations of teaching effectiveness. *Journal of Higher Education*, 64, 1-18.

Marsh, H. W., & Dunkin, M. (1992). Students' evaluations of university teaching; A multidimensional perspective. In J. C. Smart (Ed.), *Higher Education: Handbook of theory and research* (Vol. 8, pp. 143-233). New York: Agathon.

Menges, R. J., & Brinko, K. T. (1986). Effects of Student Evaluation Feedback: A Meta-Analysis of Higher Education Research. Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco: CA, April.

Stake, J. (1997). Response to Haskell: Academic Freedom, Tenure, and Student Evaluation of Faculty. *Education Policy Analysis Archives*, 5(8). Retrieved June 15, 2005 from <http://epaa.asu.edu/epaa/v5n8.html>

Theall, M. (1997). On drawing reasonable conclusions about student ratings of instruction: a reply to Haskell and to Stake. *Education Policy Analysis Archives*. Retrieved June 15, 2005 from <http://epaa.asu.edu/epaa/v5n8c2.html>.

Education Policy Analysis Archives

<http://epaa.asu.edu>

Editor: Sherman Dorn, University of South Florida

Production Assistant: Chris Murrell, Arizona State University

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Sherman Dorn, epaa-editor@shermadorn.com.

EPAA Editorial Board

Michael W. Apple
University of Wisconsin

Greg Camilli
Rutgers University

Mark E. Fetler
California Commission on Teacher
Credentialing

Richard Garlikov
Birmingham, Alabama

Thomas F. Green
Syracuse University

Craig B. Howley
Appalachia Educational Laboratory

Patricia Fey Jarvis
Seattle, Washington

Benjamin Levin
University of Manitoba

Les McLean
University of Toronto

Michele Moses
Arizona State University

Anthony G. Rud Jr.
Purdue University

Michael Scriven
Western Michigan University

Robert E. Stake
University of Illinois—UC

Terrence G. Wiley
Arizona State University

David C. Berliner
Arizona State University

Linda Darling-Hammond
Stanford University

Gustavo E. Fischman
Arizona State University

Gene V Glass
Arizona State University

Aimee Howley
Ohio University

William Hunter
University of Ontario Institute of
Technology

Daniel Kallós
Umeå University

Thomas Mauhs-Pugh
Green Mountain College

Heinrich Mintrop
University of California, Berkeley

Gary Orfield
Harvard University

Jay Paredes Scribner
University of Missouri

Lorrie A. Shepard
University of Colorado, Boulder

Kevin Welner
University of Colorado, Boulder

John Willinsky
University of British Columbia

Archivos Analíticos de Políticas Educativas

ASSOCIATE EDITORS

Gustavo E. Fischman & Pablo Gentili

Arizona State University & Universidade do Estado do Rio de Janeiro

Founding Associate Editor for Spanish Language (1998—2003)

Roberto Rodríguez Gómez

Editorial Board

Hugo Aboites

Universidad Autónoma
Metropolitana-Xochimilco

Dalila Andrade de Oliveira

Universidade Federal de Minas
Gerais, Belo Horizonte, Brasil

Alejandro Canales

Universidad Nacional Autónoma
de México

Erwin Epstein

Loyola University, Chicago,
Illinois

Rollin Kent

Universidad Autónoma de
Puebla. Puebla, México

Daniel C. Levy

University at Albany, SUNY,
Albany, New York

María Loreto Egaña

Programa Interdisciplinario de
Investigación en Educación

Grover Pango

Foro Latinoamericano de
Políticas Educativas, Perú

Angel Ignacio Pérez Gómez

Universidad de Málaga

Diana Rhoten

Social Science Research Council,
New York, New York

Susan Street

Centro de Investigaciones y
Estudios Superiores en
Antropología Social Occidente,
Guadalajara, México

Antonio Teodoro

Universidade Lusófona Lisboa,

Adrián Acosta

Universidad de Guadalajara
México

Alejandra Birgin

Ministerio de Educación,
Argentina

Ursula Casanova

Arizona State University,
Tempe, Arizona

Mariano Fernández

Enguita Universidad de
Salamanca. España

Walter Kohan

Universidade Estadual do Rio
de Janeiro, Brasil

Nilma Limo Gomes

Universidade Federal de
Minas Gerais, Belo Horizonte

Mariano Narodowski

Universidad Torcuato Di
Tella, Argentina

Vanilda Paiva

Universidade Estadual Do
Rio De Janeiro, Brasil

Mónica Pini

Universidad Nacional de San
Martín, Argentina

José Gimeno Sacristán

Universidad de Valencia,
España

Nelly P. Stromquist

University of Southern
California, Los Angeles,
California

Carlos A. Torres

UCLA

Claudio Almonacid Avila

Universidad Metropolitana de
Ciencias de la Educación, Chile

Teresa Bracho

Centro de Investigación y
Docencia Económica-CIDE

Sigfredo Chiroque

Instituto de Pedagogía Popular,
Perú

Gaudêncio Frigotto

Universidade Estadual do Rio
de Janeiro, Brasil

Roberto Leher

Universidade Estadual do Rio
de Janeiro, Brasil

Pia Lindquist Wong

California State University,
Sacramento, California

Iolanda de Oliveira

Universidade Federal
Fluminense, Brasil

Miguel Pereira

Catedrático Universidad de
Granada, España

Romualdo Portella do

Oliveira

Universidade de São Paulo

Daniel Schugurensky

Ontario Institute for Studies in
Education, Canada

Daniel Suarez

Laboratorio de Políticas
Publicas-Universidad de
Buenos Aires, Argentina

Jurjo Torres Santomé

Universidad de la Coruña,
España